Biophysical Journal Volume 103 August 2012 767–776

Hybrid Dynamics Simulation Engine for Metalloproteins

Manuel Sparta,[†] David Shirvanyants,[‡] Feng Ding,[‡] Nikolay V. Dokholyan,[‡*] and Anastassia N. Alexandrova[†*]
[†]Department of Chemistry and Biochemistry, University of California, Los Angeles, California; and [‡]Department of Biochemistry and Biophysics, University of North Carolina, Chapel Hill, North Carolina

ABSTRACT Quality computational description of metalloproteins is a great challenge due to the vast span of time- and length-scales characteristic of their existence. We present an efficient new method that allows for robust characterization of metalloproteins. It combines quantum mechanical (QM) description of the metal-containing active site, and extensive dynamics of the protein captured by discrete molecular dynamics (DMD) (QM/DMD). DMD samples the entire protein, including the backbone, and most of the active site, except for the immediate coordination region of the metal. QM operates on the part of the protein of electronic and chemical significance, which may include tens to hundreds of atoms. The breathing quantum-classical boundary provides a continuous mutual feedback between the two machineries. We test QM/DMD using the Fe-containing electron transporter protein, rubredoxin, and its three mutants as a model. QM/DMD can provide a reliable balanced description of metalloproteins' structure, dynamics, and electronic structure in a reasonable amount of time. As an illustration of QM/DMD capabilities, we then predict the structure of the Ca^{2+} form of the enzyme catechol O-methyl transferase, which, unlike the native Mg^{2+} form, is catalytically inactive. The Mg^{2+} site is ochtahedral, but the Ca^{2+} is 7-coordinate and features the misalignment of the reacting parts of the system. The change is facilitated by the backbone adjustment. QM/DMD is ideal and fast for providing this level of structural insight.

INTRODUCTION

Metalloproteins can be exceptionally hard to characterize computationally (1,2). The presence of the metal cation (or several of them) results in strong Coulombic forces acting on charged amino acids and the backbone. Indeed, proteins are known to respond dramatically to the installation or removal of metal cations, including large conformational changes, and to aggregation (3–9). More subtly, metals with incomplete populations of d-shells of atomic orbitals have strongly preferred coordination geometries, whether tetrahedral, octahedral, or square planar. This preference imposes directionality on the positions of amino acids surrounding the metal and thus impacts protein structure.

From the point of view of the metal, the structure and dynamics of the surroundings mean favorable or unfavorable accommodation of one of its preferred coordination modes. Deviations from the preferred geometry reduce the protein-metal affinity. Also, the electronic structure of a metal is highly sensitive to its coordination environment, both within and beyond the first coordination sphere. Different electronic configurations may be possible for a metal, depending on the chemical nature of its ligands; for example, Fe can be low-spin or high-spin, depending on how hard or soft, respectively, its ligands are. A very specific electronic configuration of the metal is essential for the proper biological function of a metalloprotein. Thus, the electronic structure of the metal and the macromolecular structure of the protein are intimately coupled and faithfully respond to changes in each other.

Metalloproteins operate on a variety of scales, and their functionalities are difficult to capture, predict, and design computationally. The focus of our work is the development of a balanced method for quality modeling of metalloproteins at all scales. Classical molecular dynamics (MD) or Monte Carlo (MC) modeling of metalloproteins is very challenging even in the case of structureless cations, such as Mg^{2+} or Zn^{2+}, due to the difficulties of parameterization of corresponding classical force fields (FFs). FFs can be parameterized for a metal found in specific coordination geometry and ligated by specific ligands. However, if the coordination environment changes unpredictably in the course of the simulation, for example, through ligand exchange, attachment, or dissociation, a classical description would fail. Although full quantum mechanical (QM) computational description of large biomolecules is prohibitively expensive, multiresolution methods utilize a focused approach, representing the total Hamiltonian as a sum of QM and classical (MM) components (10), with additional electrostatic (11), and boundary interaction terms. QM/MM methods aim to accurately model the active site (10–100 atoms (12,13)) using first-principles QM techniques, and yet to achieve structural and dynamic sampling of the entire system with MM, using one of the parameterized atomistic FFs, such as CHARMM (12,14,15) or AMBER (16,17). The treatment of the boundary between QM and MM regions is the most challenging part of QM/MM, especially with molecules spanning across this boundary. Typically, the QM region is capped with hydrogen atoms or electron pairs, during QM calculations. In addition, different schemes have been proposed to treat bonded and nonbonded interactions at the boundary (10,17–22).

Submitted January 12, 2012, and accepted for publication June 18, 2012.
*Correspondence: ana@chem.ucla.edu or dokh@email.unc.edu
Editor: Bert de Groot.

0006-3495/12/08/0767/10 $2.00
http://dx.doi.org/10.1016/j.bpj.2012.06.024

The QM/MM approach was first used by Levitt and Warshel (10) to calculate energies of intermediate states of enzymatic reaction. In years since the early works, QM/MM methods have undergone a remarkable advancement, making it possible to study reaction pathways in large systems, such as solvated enzymes (23,24). QM methods employed in QM/MM now range from semiempirical (25–31), to Hartree-Fock and post-Hartree-Fock methods (13,18,32,33), to density functional theory (DFT) (14,16,34–36), as well as ab initio molecular dynamics (AIMD) (17,37–40) and Car-Parrinello MD (37,41). Car-Parrinello-MD-based algorithms are especially reliable, and the spectacular advances made in their development over recent years have improved both performance and parallelization efficiency (42,43). However, they are still quite expensive, and simulations beyond nanosecond-range and large-scale structure optimizations of large protein systems are out of reach. The problem of affordability is so far best solved by running dynamic QM/MM simulations on a graphics processing unit (44–46), but QM codes for graphics processing units are not fully developed yet. Various advanced methods have been proposed for calculations of electrostatic interactions between the QM and MM regions (11,17,18,32,38,40). At the same time, the MM component has been more conservative, commonly being variants of classical MD (14,35,47) or MC (25–31,34).

Modern QM/MM applications can be loosely categorized (47) into analysis of specific system states and reaction intermediates (12,16,18,32,48–51), problems requiring local geometry optimization of known structures (13–15,33,38,52,53), or full energy landscape sampling for global minimization (13,34–36). Current QM/MM methods are most instrumental when a high-resolution starting structure of the protein is available. However, to date, no QM/MM method can robustly predict the dynamic response of a nonmetal protein to the installation of a metal. Similarly, no method can dock substrates or inhibitors to metalloproteins, if significant protein adjustment is required to mold around the substrate. In general, there is no method of performing affordable extensive sampling of metalloproteins for structural or mechanistic studies.

Capitalizing on the previously demonstrated ability of discrete MD (DMD) (54–57) to extensively sample protein conformations, we have designed the hybrid QM/DMD approach to study protein systems, undergoing large conformational rearrangements. In QM/DMD, classical MD is replaced by the recently developed method of atomistic event-driven DMD (54–57). In DMD, equations of motion are integrated along space coordinates, and not time, as in classical MD. This feature allows easier conjugation of QM and DMD regions, since DMD permits discontinuous potentials and does not need to compute or input interaction forces. In QM/DMD, the management areas of QM and DMD in the protein structure overlap. Via this shared-domain approach, structural information can be exchanged between the QM and MM (DMD) regions. This approach makes it possible to model complex active sites, which are not directly supported by classical FFs. For an additional acceleration, QM/DMD takes advantage of partial geometry relaxations on the Born-Oppenheimer potential energy surface of the metal-containing active site, instead of quantum dynamics. Finally, DMD algorithm and FFs have been optimized for efficient sampling of protein chains, making the combined QM/DMD method best suitable for studies of metalloproteins and enzymes.

We test our method on the electron-transporting iron-sulfur protein, rubredoxin. Present in numerous forms of life, it exhibits backbone structural similarity and 50–60% sequence similarity among homologs. However, intriguingly, it covers a wide range of sequence-dependent reduction potentials, between −80 and +40 mV (58), which found some rationalization in recent theoretical work by Knapp et al. (59,60). Thus, rubredoxin and its mutants represent a simple and yet intricate model for revealing the strengths and limitations of our, to our knowledge, new method at a range of scales. Specifically, we predict the macro- and microscopic structures of rubredoxin and its mutants, restore distorted rubredoxin to the equilibrium geometry, and predict the changes in reduction potential in response to mutations in the second coordination sphere of Fe.

We further highlight the capabilities of QM/DMD by predicting the structural effects of the metal substitution by modeling ion exchange in the catechol-O-methyl transferase (COMT) enzyme. COMT is an enzyme involved in the biology of pain; it degrades catecholamines (such as dopamine) by methylating them in the presence of a divalent metal cation (Mg^{2+}) and S-adenosyl-L-methione (SAM). The role of the Mg^{2+} ion is merely coordination and positioning of the substrate (61). Yet surprisingly, the Ca^{2+} form of COMT is catalytically inactive. In many cases, Mg^{2+} and Ca^{2+} act as physiological antagonists of each other, but in the case of COMT, the mechanism of its inhibition by Ca^{2+} remains unclear.

QM/DMD METHOD

Partitioning the protein

One of the key features of QM/DMD is its approach to partitioning the protein into the areas of management of QM and DMD (Fig. 1 *A*). Our strategy is to sample with DMD all of the protein but the region immediately surrounding the metal center (Fig. 1 *A*, *dark gray*). On the other hand, the chemically meaningful description of the active site, needed to assess the electronic structure of the metal, or simulating mechanisms of catalyzed reactions, requires a QM treatment of a fairly large area in the protein, including the metal, its immediate ligands, substrates, and occasionally amino acids of the second coordination sphere.

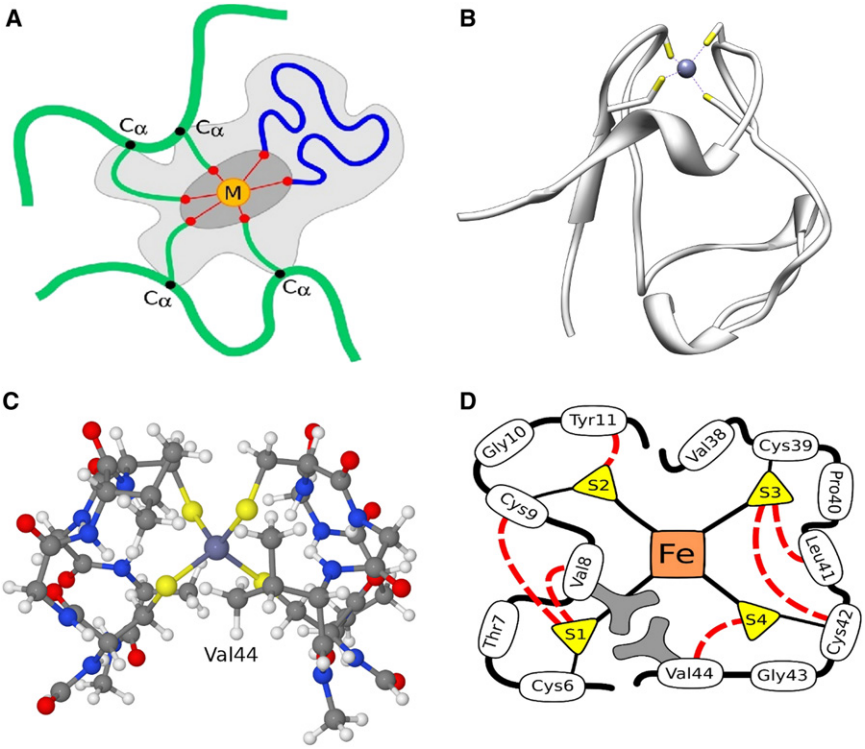

FIGURE 1 (*A*) Areas of management of QM and DMD in QM/DMD simulations. The metal is denoted M. Only QM can move the dark gray area. The light gray area is managed by both QM and DMD, depending on the phase of the simulation. It contains M, its ligands, such as amino acids up to Cα (*thin green lines*) and substrates (*blue line*), and potentially other amino acids and cofactors. The rest of the system is sampled by DMD only. (*B*) The full rubredoxin protein. (*C*) The QM-DMD region used in QM/DMD simulations of rubredoxin. (*D*) Schematic representation of the QM-DMD region: black lines indicate covalent bonding; red dashed lines indicate H-bonds between the S_{Cys} and the backbone; side chains of Val-8 and Val-44 are light gray.

The region treated at the QM level in QM/DMD contains hundreds of atoms (Fig. 1 *A*, *light gray*). This shared QM-DMD region is modified quantum mechanically through gradient after relaxation, instead of ab initio dynamics. Thus, there are a total of three regions in the protein: QM only (*dark gray*), QM-DMD (*light gray*), and DMD only (Fig. 1 *A*). The QM-DMD boundary is breathing, and goes around the dark gray region or the light gray region, depending on the stage in the simulation. A breathing boundary ensures extensive communication between the regions, as will be demonstrated. Hence, besides capping with H atoms, no further treatment of the boundaries, such as electrostatic embedding, is introduced. Indeed, even traditional QM/MM with stationary boundaries, performed with and without embedding schemes, has been shown to produce comparable errors, questioning the quality of and need for any embedding (62).

QM and DMD constituents

In the simulations reported here, DFT was used in the description of the QM region, as it has become the workhorse for simulation of biologically relevant metals (63). The BP86 (64–68) functional was used in conjunction with the resolution of identity and multipole accelerated resolution of identity (69,70), empirical dispersion correction (71), and the double zeta quality basis set (def2-SV(P)) (72). In evaluations of reduction potentials, single-point energy calculations were done with TPPS (73,74)/def2-TZVPP (75). To account for the effect of the solvent, the conductorlike screening model (COSMO) continuum solvation model was used (76). The dielectric constant of the implicit solvent was set to 78.4, since the QM-DMD region is mostly surrounded by water rather than the protein. TURBOMOLE v.6.3 (77) was used for all QM calculations.

DMD was described in detail elsewhere (54–57). Briefly, it is the type of MD that is based on solving the equations of conservation of momentum and energy, rather than Newtonian equations of motion. DMD uses the implicit solvent atomistic model of proteins, based on the discretized CHARMM/EEF1 FF (78). It operates on a set of square-well potentials, such that interactions between atoms are nonzero only when atoms are within the given square well. DMD was proven to be exceptionally fast and efficient in sampling of biomolecules for a variety of applications. Here, DMD underwent several modifications to become more context-dependent in consistency with QM (Supporting Material).

Simulation strategy

Simulations are iterative. First, DMD is run on all but the QM/only region, which remains fixed. A frozen QM/only region during the DMD phase means that the DMD engine is virtually unaware of the exact nature of the metal but responds to the metal + surrounding atoms moiety. This in turn means that any metal or active site can be simulated, even if its DMD FF parameterization is not available. In

addition, atoms in the shared QM-DMD region have constraints imposed on them to maintain bond lengths and angles relative to the frozen QM region in accord with those observed during the QM phase (see Section Ib in the Supporting Material). All other interactions between the DMD, QM, and QM-DMD regions (such as Van der Waals, solvation, and H-bond formation) are calculated in the same way as interactions within the DMD region. The simulation proceeds in an annealing regime for the first 3,500 DMD time units (t.u.) and then at a low temperature for 10,000 t.u. (0.5 ns), at which point data are collected. The cold ensemble of 10,000 structures is clustered according to geometric similarity. Geometric centroids and lowest-energy structures are selected from each cluster, and single-point QM energies of their QM-DMD regions are calculated. Structures are then evaluated, based on the combined scoring index (*SI*), which depends on both the QM and DMD energies and their differences from the minimum energies found for the given set of structures:

$$SI^{i} = n\left(E^{i}_{DMD} - E^{min}_{DMD}\right) + (1-n)\left(E^{i}_{QM} - E^{min}_{QM}\right)$$

In this work, n is set to 0.5. The structure having the smallest *SI* is chosen, and its QM-DMD region is partially optimized quantum mechanically, with fixed positions for the atoms bridging the QM region to the rest of the protein. The QM optimization is bottleneck for the procedure, and to accelerate the method, we found it adequate to interrupt the optimizations after a preset number of steps (100–200), when optimization is typically close to convergence. However, tightly optimized structures are still obtained to calculate the properties of the metal. The partially optimized QM region is then reinstalled in the protein, and the QM-DMD boundary shrinks back from the dark gray region. In the course of the following iteration, in the DMD stage, the structural information from the optimized QM-DMD region can propagate to the rest of the protein. The alternating QM and DMD jobs with the breathing QM-DMD boundary are performed for 20 to hundreds of iterations, depending on the task (this roughly corresponds to 10–50 ns of dynamics). More specific details of QM/DMD can be found in the Supporting Material.

CAPABILITIES AND LIMITATIONS OF QM/DMD

We test the performance of QM/DMD on the electron-transporting iron-sulfur protein rubredoxin (PDB ID 1FHM (79)) and its mutants. The protein consists of 52 residues (Fig. 1 *B*), 13 of which are included in the QM-DMD region (Fig. 1 *C*). This QM-DMD region is well structured and possesses numerous H-bonds in which the Fe-bound deprotonated S_{Cys} atoms serve as H-bond acceptors (80,81). The major H-bond donors are the six backbone amide groups (Fig. 1 *D*). In addition, there are several unusual H-bonds where the donors are the aliphatic C-H groups of Val, which normally do not engage in H-bonding (81). We test QM/DMD at three scales: large (structure recapitulation and recovery), intermediate (subtle structural details, such as lengths of H-bonds), and small (electronic properties of the metal center with the accuracy of 1–2 kcal/mol from the experimental data).

Large-scale test 1: structure recapitulation

Our first task is to test stability of the QM/DMD simulated wild-type (WT) rubredoxin. The simulations starting from the x-ray structure lead the protein away from it, stabilizing in the found QM/DMD basin (Fig. 2 *A*). The RMSD of the entire protein backbone between the QM/DMD and the x-ray structures stabilizes at 1.4 Å, and the all-atom RMSD of the QM-DMD region stabilizes at 1.1 Å (Fig. 2 *A*). The QM energy of the QM-DMD region is conserved to -6320.387 ± 0.005 a.u. for the reduced form of the protein (6.3 ± 3.0 kcal/mol above the QM energy of the QM-DMD region obtained in the first iteration). The DMD energy also converges, though fluctuations are larger than those of the QM energies due to a larger number of atoms in the DMD region.

Similar robust performance of QM/DMD was found for the three mutants of rubredoxin, V44I, V44A, and V44G (Fig. 3). The structure of the V44A variant agrees well with the available x-ray structure (PDB ID 1C09 (82); see Fig. S10 in the Supporting Material). For the V44G mutant, we observe a fairly large void space in the area of the missing side chain of residue 44 (Fig. 3 *C*). This area is likely to be solvent-accessible. An interesting structural feature was detected for the V44I mutant (Fig. 3, *A* and *B*). The bulkier side chain of Ile does not fit well within the space in the protein originally occupied by Val, so adjustments and repacking of the protein were required to accommodate Ile-44. QM/DMD predicts the existence of two isoenergetic

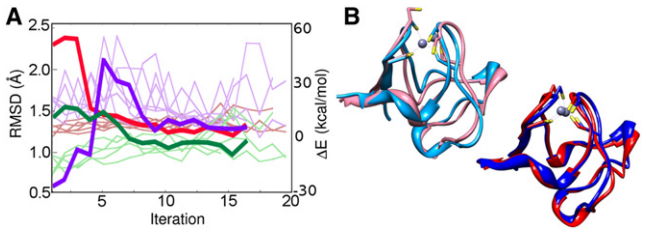

FIGURE 2 Large-scale testing of QM/DMD: structure recapitulation and recovery for reduced rubredoxin. (*A*) The QM-DMD region RMSD from the x-ray structure (*light green lines*), the QM energies (*pink lines*), and the DMD energies (*light purple lines*) stabilize by the fifth iteration. Thick lines illustrate the return of the distorted WT structure to equilibrium in terms of the QM-DMD region RMSD (*green*), QM energy (*red*), and DMD energy (*purple*). (*B*) The overlay of the x-ray structure (*light blue*) and the distorted structure (*pink*), and of representative QM-DMD equilibrated structures, starting from x-ray (*dark blue*) and from the distorted structure (*red*).

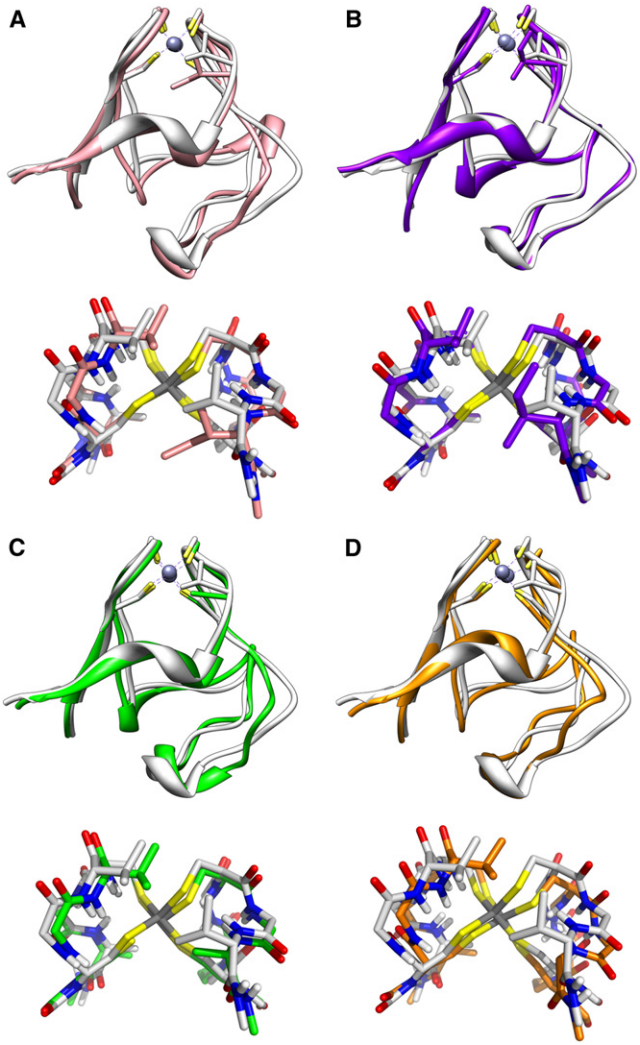

FIGURE 3 Overlay of the x-ray structure of the WT rubredoxin (*white*) and representative QM/DMD snapshots of equilibrated structures of the three mutants: two isoenergetic variants of V44I, with I44 pointing down (*A*) and up (*B*), V44A (*C*), and V44G (*D*). The side chains are shown for the four Cys residues and residue 44. The overlays of the QM-DMD regions are presented beneath each structure. Nonpolar H atoms are omitted for clarity.

structures for V44I. In one, Ile-44 points up toward the Fe center, and in the other it points down and away from the metal. All structures are available upon request.

Large-scale test 2: structure recovery from deformation

Next, we demonstrate that QM/DMD can restore a distorted structure of rubredoxin, which also is a test for the general ability of QM/DMD to predict the structural response of a protein to the installation of a metal. We also show that metalloprotein structure stabilization is a result of cooperation between QM and DMD and is not due to the specific traits of either component.

First, rubredoxin was distorted from the equilibrium by removing Fe, protonating the four Cys residues coordinating Fe, and allowing DMD to relax the protein for 10,000 t.u. Not surprisingly, even in this short relaxation, the demetallated structure visibly changed, and the DMD energy of the protein decreased. Fe was then reinstalled into the distorted protein, and the four Cys residues were once again deprotonated. In the distorted structure, the tetrahedral coordination of the Fe center was disrupted, and one Cys was completely detached (Fig. 2 *B*). This structure was then passed to QM/DMD for reequilibration.

The starting QM energy of the distorted QM-DMD region is higher than that of the WT by ~50 kcal/mol. The starting DMD energy is lower than that of the WT by ~40 kcal/mol. The QM/DMD simulation returns the protein at the original WT basin, approximately by the eighth iteration (Fig. 2 *A*). During the relaxation process, the conformation of the entire protein changes to accommodate the metal center. The Cys residues return to coordinating Fe. The actual event of Cys-Fe binding is controlled purely at the QM level. Since in this case it was known that the four Cys residues must eventually bind Fe, the S-Fe moiety constituted the QM-only region even at large S-Fe separations. The RMSD values converge to those characteristic of the WT. The QM energy decreases, the DMD energy increases accordingly, and both also converge to the values from the simulation started from the x-ray structure. This result shows that QM/DMD energy function has at least a local minimum at the native structure. It also illustrates the quality of the communication between QM and DMD in QM/DMD and the chosen *SI* function.

A more challenging problem arises in the case of ion substitution. We substituted Fe(II) in rubredoxin with Pd(II) in a low-spin state, which is known to prefer a square planar coordination. QM/DMD simulation yields a structure of the protein representing a clear compromise between the structural preferences of Pd (square planar) and the rest of the protein (accommodating a tetrahedrally coordinated metal). The protein is distorted, and the metal coordination sphere approached, though it did not fully reach, square planar geometry. Details of this test are presented in the Supporting Material. Overall, we conclude that QM/DMD is capable of reproducing the protein structure and recovering it after a moderate distortion from the equilibrium.

Intermediate-scale test: finer structural details sensitive to the oxidation state of Fe and mutations

Next, we test that QM/DMD reproduces structural trends in the QM-DMD region with chemical accuracy. Upon oxidation from Fe^{2+} to Fe^{3+}, the Fe coordination is expected to become tighter due to the stronger coulombic attraction of the ligands to the cation. This is observed experimentally: the average Fe-S_{Cys} distance in the Fe^{2+}

form (PDB ID 1FHM) is 2.36 ± 0.03 Å, and that in the Fe^{3+} form (PDB ID 1FHH) is 2.27 ± 0.02 Å. This trend is reproduced in our simulations. The calculated Fe-S_{Cys} distances are 2.32 ± 0.03 Å and 2.21 ±0.04 Å, for the Fe^{2+} and Fe^{3+} forms, respectively. At the same time, the six H-bonds formed by S_{Cys} atoms ligating Fe and the backbone N-H groups become weaker and longer on average by 5–18%, as expected (Table S1).

Next, we focus on the specific H-bond between S_{Cys42} and the backbone N-H group of Val-44 in the WT. According to Lin et al. (80), the reduction potential of Fe can be modulated by mutations of Val at this position, which is quite remarkable, considering that residue 44 belongs to the second coordination sphere of Fe. It was hypothesized that this impact was due to the changing strength and length of the H-bond of Cys-42 to the backbone of residue 44. However, at present, not all of the structures of the mutants are available to fully confirm this hypothesis, and the H atoms are invisible for the x-ray diffraction. According to QM/DMD, of all H-bonds in the QM-DMD region, the H-bond to the backbone of the residue 44 undergoes the strongest change upon mutation (10–18%). In full accord with experimentally derived H-bond distances, it gets progressively shorter in the sequence WT, V44I, V44A, and V44G. In the Fe^{3+} form, its length is 3.07, 3.03/2.77 (depending on the conformation of Ile-44), 2.55, and 2.50 Å, and in the Fe^{2+} form it is 2.63, 2.75/2.49, 2.33, and 2.29 Å, respectively (Fig. 4). This confirms that QM/DMD reliably reproduces these rather delicate structural details.

Small-scale test: reduction potential as a function of off-site mutations

We used the QM/DMD ensembles of structures to estimate the changes in reduction potential in the mutants and compare those to the experiment. The reduction potential, E_0, changes from −77 mV in the WT to −53 mV in V44I, −24 mV in V44A, and 0 mV in V44G (80). Through the definition of E_0, these values translate to ΔG of Fe reduction, and its changes ($\Delta\Delta G$) upon mutation, which read as 0.0 kcal/mol (reference, WT), −0.6 kcal/mol (V44I), −1.2 kcal/mol (V44A), and −1.8 kcal/mol (V44G). Reproducing the ΔG and $\Delta\Delta G$ trends is a challenging test for QM/DMD, calling for true chemical accuracy from the method. In fact, this accuracy would be on the verge of what modern high levels of ab initio theory can do, and so a qualitative agreement with the experiment would be adequate for QM/DMD.

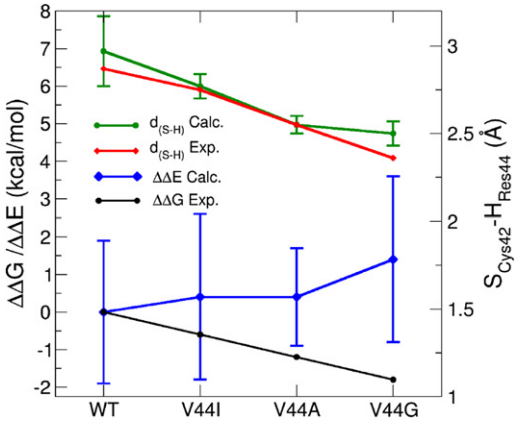

FIGURE 4 Experimental $\Delta\Delta G$ and predicted $\Delta\Delta E$ values for the four variants of rubredoxin, and the corresponding theoretical and experimentally derived H-bond lengths between S_{Cys42} and the backbone of residue 44 for the oxidized form of the protein.

We make two assumptions in our estimations of $\Delta\Delta G$. First, the experimentally measured E_0 is an adiabatic quantity, i.e., the protein has a chance to fully relax in the time frame of the measurement, and the E_0 corresponds to ΔG between the minima of the Fe^{2+} and Fe^{3+} forms of the protein. The second assumption (based on Marcus theory (83) and our QM/DMD simulations for the WT) is that the overall structure and dynamics of the protein do not change substantially upon oxidation/reduction, except for the slight tightening of the Fe coordination sphere. Thus, the associated change of entropy, ΔS, is small and is similar for all the mutants, and $\Delta\Delta S$, characterizing the difference in the change of entropy for different mutants, can be ignored in predicting $\Delta\Delta G$. Hence, $\Delta\Delta G$ can be approximated by QM $\Delta\Delta E$.

ΔE and $\Delta\Delta E$ values averaged over QM/DMD ensembles were calculated with inclusion of implicit solvation. Without solvation, experimental data could not be qualitatively reproduced, especially for the V44G mutant. From the QM/DMD structure of V44G, we know that there is a surface-accessible void space beneath S_{Cys42}. Hence, it is likely that solvation of this area is partially responsible for the change in the reduction potential of V44G. Calculated QM ΔE and $\Delta\Delta E$ are qualitatively correct, and for all four proteins the ΔE values fall into a narrow range, in agreement with the experimental spread of ΔG values (Table S2). In Fig. 4, $\Delta\Delta E$ values are compared to the experimental $\Delta\Delta G$ values. The predicted values of the $\Delta\Delta E$ for each protein fall within an ~4-kcal/mol window due to the structural fluctuations in each ensemble. Qualitatively, the experimentally observed trend for $\Delta\Delta G$ and the theoretical trend for $\Delta\Delta E$ agree, which is quite remarkable considering the resolution of the theoretical method. Indeed, here, we aim to resolve differences in $\Delta G < 2$ kcal/mol. Even for high-level ab initio calculations on much smaller systems than the QM-DMD region chosen for rubredoxin, this level of resolution would be nearly prohibitive. For example, in predictions of photoelectron spectra of clusters smaller than 10 atoms, using methods including a large amount of electron correlation, an accuracy of 0.1–0.2 eV (2–5 kcal/mol) is considered normal and sufficient (84,85). Thus, the accuracy of the QM/DMD method is

virtually the same as that of regular stationary DFT for systems of the same size as our QM-DMD region or smaller. Others have reported more sophisticated calculations of the reduction potentials of rubredoxin that produce a better agreement with experiment (59,60). Such quantitative techniques can be applied to the obtained QM/DMD ensemble, but such elaboration is not the goal of this study. Overall, we conclude that QM/DMD is capable of semi-quantitative description of the delicate electronic effects at the metal center.

The active Mg^{2+} and inactive Ca^{2+} forms of COMT

Replacement of the metal ion can have dramatic effects on protein structure, especially when the two ions have different coordination. An example of such ion-driven structural rearrangement is the enzyme-inactivating replacement of the native Mg^{2+} with Ca^{2+} in the active site of COMT.

Mg^{2+} in COMT is coordinated octahedrally by the side chains of Asp-141, Asp-169, and Asn-170, one water molecule, and the substrate (61). In QM/DMD simulations, the shared QM-DMD region included all these moieties as well as the side chains of Lys-144 and Glu-199 that form H-bonds to the substrate and the side chain of Met-40 that shields the active site from solvent. Using QM/DMD, we recapitulated the native Mg^{2+} structure of COMT. As compared to the available x-ray structure (61), the average heavy-atom RMSD of the QM-DMD region is 0.7 Å and that of the backbone of the protein is 1.2 Å. A representative structure of the active site is shown in Fig. 5 *A*. SAM and the O atom of catechol, which act as the methyl group donor and acceptor, are aligned for transfer. Substitution of Mg^{2+} by Ca^{2+} induces a significant change in the coordination environment of the metal: contrary to Mg^{2+}, Ca^{2+} is coordinated by both oxygen atoms of Asp-141 and is thus 7-coordinate. In this configuration, Asp-141 is unable to engage in an H-bond with the water molecule, which instead interacts with Glu-199 (Fig. 5 *B*). Finally, the catechol O atom and the methyl group of SAM are not properly aligned for methyl transfer. Changes are accompanied by the adjustment of the backbone (backbone RMSD from the crystal structure is 1.3 Å). It is important to stress that the observed repacking of the active site could not be obtained by a simple substitution of the metal in the crystal structure and subsequent geometry optimization of the active site at the fixed position of the backbone (Fig. S15). Analysis of Ca^{2+} COMT may shed light on the mechanism of COMT inhibition.

CONCLUSIONS

Modeling metalloproteins is a challenge for computational biology due to the strong interactions and coupling between the ion and protein conformations. We reported on a new, to our knowledge, method for modeling metalloproteins, QM/DMD. It combines the capabilities of QM—producing ab initio-quality energies, genuine wave-function-driven structures, and electronic properties of the metal-containing active sites—with those of DMD—aggressively sampling the protein, including large-scale domain motions. Innovative algorithms, such as a breathing QM-DMD boundary, are implemented in QM/DMD to make simulations on experimentally relevant scales computationally affordable. QM/DMD also takes advantage of more traditional techniques, such as annealing, clustering, and scoring based on both QM and DMD. QM/DMD is successful in modeling metalloproteins at a range of scales, from large (>10 Å), to intermediate (1–10 Å), to small (<1 Å).

We tested QM/DMD on the Fe-S electron-transporting protein rubredoxin and several of its mutants. QM/DMD faithfully reproduces the structures of the metalloproteins, in agreement with x-ray data, and even down to the finest structural details, such as lengths of weak H-bonds and their variations upon mutations in the protein. The method also can restore the structure of the protein to equilibrium after a mild distortion due to the property of the combined potential energy function reaching its minimum at the native structure. With the structural information provided by QM/DMD, delicate electronic properties of the active site, such as changes in the reduction potential of the metal upon off-active-site mutations, can be qualitatively reproduced. In view of the subtlety of these changes (on the order of tens of mV, i.e., ~k_BT), and the large size of QM-treated active sites, even a qualitative agreement is quite remarkable.

We illustrated the capabilities of the method on the Mg^{2+} and Ca^{2+} forms of COMT. Metal replacement in this protein is found to lead to the active-site repacking and adjustment of the protein backbone, which could not be captured

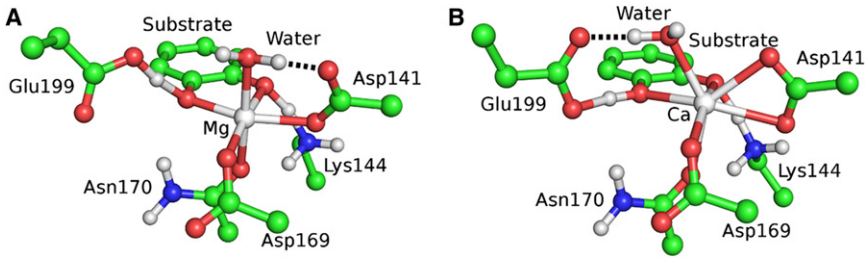

FIGURE 5 Coordination of the ion in the active site of COMT depends on the nature of the metal. (*A*) Mg^{2+}. (*B*) Ca^{2+}. Nonpolar hydrogens are removed for clarity.

without a combination of QM treatment of the active site and sampling of the protein structure. QM/DMD is a general method that can be used for sampling structures of metalloproteins, predicting conformational changes in proteins upon metal installation or replacement, and changes in the electronic structure of the metal in response to the changes in the protein, and flexible docking of substrates and inhibitors to metal-containing active sites. In particular, it can be instrumental for structural and mechanistic studies on metalloenzymes. Developed software is available upon request; its use is subject to the license agreements of the parallel version of DMD (available for distribution through Molecules in Action, LLC, http://moleculesinaction.com), and TURBOMOLE (77).

SUPPORTING MATERIAL

Details of the QM/DMD method, computational details, additional results, 15 figures, and references (86–90) are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(12)00681-9.

This work was supported by the DARPA Young Faculty Award N66001-11-1-4138 (A.N.A.), the University of California through the Faculty Career Development Award and Faculty Research Grant (A.N.A.), and National Institutes of Health grant R01GM080742 (N.V.D.).

REFERENCES

1. Chung, L. W., X. Li, and K. Morokuma. 2010. Modeling enzymatic reactions in metalloenzymes and photobiology by quantum mechanics (QM) and quantum mechanics/molecular mechanics (QM/MM) calculations. *In* Quantum Biochemistry. C. F. Matta, editor. Wiley-VCH, Weinheim, Germany. 85–128.
2. Warshel, A. 1997. Computer Modeling of Chemical Reactions in Enzymes and Solutions. John Wiley & Sons, New York.
3. Keating, K. M., L. R. Ghosaini, …, J. M. Sturtevant. 1988. Thermal denaturation of T4 gene 32 protein: effects of zinc removal and substitution. *Biochemistry.* 27:5240–5245.
4. Matt, T., M. A. Martinez-Yamout, …, P. E. Wright. 2004. The CBP/p300 TAZ1 domain in its native state is not a binding partner of MDM2. *Biochem. J.* 381:685–691.
5. Ding, F., and N. V. Dokholyan. 2008. Dynamical roles of metal ions and the disulfide bond in Cu, Zn superoxide dismutase folding and aggregation. *Proc. Natl. Acad. Sci. USA.* 105:19696–19701.
6. Banci, L., I. Bertini, …, M. S. Viezzoli. 2003. Solution structure of Apo Cu,Zn superoxide dismutase: role of metal ions in protein folding. *Biochemistry.* 42:9543–9553.
7. Cao, Y., T. Yoo, and H. Li. 2008. Single molecule force spectroscopy reveals engineered metal chelation is a general approach to enhance mechanical stability of proteins. *Proc. Natl. Acad. Sci. USA.* 105:11152–11157.
8. Blum, O., A. Haiek, …, H. B. Gray. 1998. Isolation of a myoglobin molten globule by selective cobalt(III)-induced unfolding. *Proc. Natl. Acad. Sci. USA.* 95:6659–6662.
9. Radford, R. J., P. C. Nguyen, …, F. A. Tezcan. 2010. Controlled protein dimerization through hybrid coordination motifs. *Inorg. Chem.* 49:4362–4369.
10. Warshel, A., and M. Levitt. 1976. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* 103:227–249.
11. Bakowies, D., and W. Thiel. 1996. Semiempirical treatment of electrostatic potentials and partial charges in combined quantum mechanical and molecular mechanical approaches. *J. Comput. Chem.* 17:87–108.
12. Svensson, M., S. Humbel, …, K. Morokuma. 1996. ONIOM: A multilayered integrated MO + MM method for geometry optimizations and single point energy predictions. A test for Diels-Alder reactions and Pt(P(t-Bu)3)2 + H2 oxidative addition. *J. Phys. Chem.* 100:19357–19363.
13. Field, M. J., P. A. Bash, and M. Karplus. 1990. A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J. Comput. Chem.* 11:700–733.
14. Eichinger, M., P. Tavan, …, M. Parrinello. 1999. A hybrid method for solutes in complex solvents: density functional theory with empirical force fields. *J. Chem. Phys.* 110:10452–10467.
15. Das, D., K. P. Eurenius, …, B. R. Brooks. 2002. Optimization of quantum mechanical molecular mechanical partitioning schemes: Gaussian delocalization of molecular mechanical charges and the double link atom method. *J. Chem. Phys.* 117:10534–10547.
16. Stanton, R. V., D. S. Hartsough, and K. M. Merz. 1993. Calculation of solvation free energies using a density functional/molecular dynamics coupled potential. *J. Phys. Chem.* 97:11868–11870.
17. Vreven, T., K. S. Byun, …, M. J. Frisch. 2006. Combining quantum mechanics methods with molecular mechanics methods in ONIOM. *J. Chem. Theory Comput.* 2:815–826.
18. Singh, U. C., and P. A. Kollman. 1986. A combined ab initio quantum mechanical and molecular mechanical method for carrying out simulations on complex molecular systems: applications to the CH3Cl + Cl^- exchange reaction and gas phase protonation of polyethers. *J. Comput. Chem.* 7:718–730.
19. Thery, V., D. Rinaldi, …, G. G. Ferenczy. 1994. Quantum mechanical computations on very large molecular systems: the local self-consistent field method. *J. Comput. Chem.* 15:269–282.
20. Maseras, F., and K. Morokuma. 1995. IMOMM—a new integrated ab initio plus molecular mechanics geometry optimization scheme of equilibrium structures and transition states. *J. Comput. Chem.* 16:1170–1179.
21. Zhang, Y. K., T. S. Lee, and W. T. Yang. 1999. A pseudobond approach to combining quantum mechanical and molecular mechanical methods. *J. Chem. Phys.* 110:46–54.
22. Philipp, D. M., and R. A. Friesner. 1999. Mixed ab initio QM/MM modeling using frozen orbitals and tests with alanine dipeptide and tetrapeptide. *J. Comput. Chem.* 20:1468–1494.
23. Hu, H., and W. T. Yang. 2008. Free energies of chemical reactions in solution and in enzymes with ab initio quantum mechanics/molecular mechanics methods. *Annu. Rev. Phys. Chem.* 59:573–601.
24. Kamerlin, S. C. L., S. Vicatos, …, A. Warshel. 2011. Coarse-grained (multiscale) simulations in studies of biophysical and chemical systems. *Annu. Rev. Phys. Chem.* 62:41–64.
25. Guimarães, C. R. W., M. Udier-Blagović, and W. L. Jorgensen. 2005. Macrophomate synthase: QM/MM simulations address the Diels-Alder versus Michael-Aldol reaction mechanism. *J. Am. Chem. Soc.* 127:3577–3588.
26. Tubert-Brohman, I., O. Acevedo, and W. L. Jorgensen. 2006. Elucidation of hydrolysis mechanisms for fatty acid amide hydrolase and its Lys142Ala variant via QM/MM simulations. *J. Am. Chem. Soc.* 128:16904–16913.
27. Alexandrova, A. N., D. Röthlisberger, …, W. L. Jorgensen. 2008. Catalytic mechanism and performance of computationally designed enzymes for Kemp elimination. *J. Am. Chem. Soc.* 130:15907–15915.
28. Kaminski, G. A., and W. L. Jorgensen. 1998. A QM/MM method based on CM1A charges: applications to solvent effects on organic equilibria and reactions. *J. Phys. Chem. B.* 102:1787–1796.
29. Chandrasekhar, J., S. Shariffskul, and W. L. Jorgensen. 2002. QM/MM simulations of cycloaddition reactions in water: contribution of enhanced hydrogen nonding at the transition state to the solvent effects. *J. Phys. Chem. B.* 106:8078–8085.

30. Acevedo, O., and W. L. Jorgensen. 2004. Solvent effects and mechanism for a nucleophilic aromatic substitution from QM/MM simulations. *Org. Lett.* 6:2881–2884.

31. Gao, J. L. 1992. Absolute free energy of solvation from Monte Carlo simulations using combined quantum and molecular mechanical potentials. *J. Phys. Chem.* 96:537–540.

32. Luzhkov, V., and A. Warshel. 1991. Microscopic calculations of solvent effects on absorption spectra of conjugated molecules. *J. Am. Chem. Soc.* 113:4491–4499.

33. Thompson, M. A., E. D. Glendening, and D. Feller. 1994. The nature of K^+ crown-ether interactions—a hybrid quantum mechanical-molecular mechanical study. *J. Phys. Chem.* 98:10465–10476.

34. Tunon, I., M. T. C. Martins-Costa, …, J. L. Rivail. 1996. A coupled density functional-molecular mechanics Monte Carlo simulation method. The water molecule in liquid water. *J. Comput. Chem.* 17:19–29.

35. Tunon, I., M. T. C. Martins-Costa, …, M. F. Ruiz-Lopez. 1997. Molecular dynamics simulations of elementary chemical processes in liquid water using density functional and molecular mechanics potentials. I. Proton transfer in strongly H-bonded complexes. *J. Chem. Phys.* 106:3633–3642.

36. Strand, M., M. T. C. Martins-Costa, …, J. L. Rivail. 1997. Molecular dynamics simulations of elementary chemical processes in liquid water using density functional and molecular mechanics potentials. II. Charge separation processes. *J. Chem. Phys.* 106:3643–3657.

37. Car, R., and M. Parrinello. 1985. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.* 55:2471–2474.

38. Biswas, P. K., and V. Gogonea. 2005. A regularized and renormalized electrostatic coupling Hamiltonian for hybrid quantum-mechanical-molecular-mechanical calculations. *J. Chem. Phys.* 123:164114–164119.

39. Carloni, P., U. Rothlisberger, and M. Parrinello. 2002. The role and perspective of ab initio molecular dynamics in the study of biological systems. *Acc. Chem. Res.* 35:455–464.

40. Laio, A., J. VandeVondele, and U. Rothlisberger. 2002. A Hamiltonian electrostatic coupling scheme for hybrid Car-Parrinello simulations. *J. Chem. Phys.* 116:6941–6947.

41. Iftimie, R., P. Minary, and M. E. Tuckerman. 2005. Ab initio molecular dynamics: concepts, recent developments, and future trends. *Proc. Natl. Acad. Sci. USA.* 102:6654–6659.

42. Bylaska, E. J., K. Tsemekhman, …, H. Jonsson. 2011. Parallel implementation of Γ-point pseudopotential plane-wave DFT with exact exchange. *J. Comput. Chem.* 32:54–69.

43. Hutter, J., and A. Curioni. 2005. Car-Parrinello molecular dynamics on massively parallel computers. *ChemPhysChem.* 6:1788–1793.

44. Ufimtsev, I. S., and T. J. Martínez. 2008. Quantum chemistry on graphical processing units. 1. Strategies for two-electron integral evaluation. *J. Chem. Theory Comput.* 4:222–231.

45. Ufimtsev, I. S., and T. J. Martínez. 2009. Quantum chemistry on graphical processing units. 2. Direct self-consistent field implementation. *J. Chem. Theory Comput.* 5:1004–1015.

46. Ufimtsev, I. S., and T. J. Martínez. 2009. Quantum chemistry on graphical processing units. 3. Analytical energy gradients and first principle molecular dynamics. *J. Chem. Theory Comput.* 5:2619–2628.

47. Senn, H. M., and W. Thiel. 2007. Atomistic approaches in modern biology. *In* Quantum Chemistry to Molecular Simulations. M. Reiher and L. Bertini, editors. Springer, Berlin. 173–290.

48. Reddy, M. R., and M. D. Erion. 2001. Calculation of relative binding free energy differences for fructose 1,6-bisphosphatase inhibitors using the thermodynamic cycle perturbation approach. *J. Am. Chem. Soc.* 123:6246–6252.

49. Thompson, M. A., and G. K. Schenter. 1995. Excited states of the bacteriochlorophyll-b dimer of *Rhodopseudomonas viridis*—a QM/MM study of the photosynthetic reaction center that includes MM polarization. *J. Phys. Chem.* 99:6374–6386.

50. Alagona, G., P. Desmeules, …, P. A. Kollman. 1984. Quantum mechanical and molecular mechanical studies on a model for the dihydroxyacetone phosphate-glyceraldehyde phosphate isomerization catalyzed by triose phosphate isomerase (TIM). *J. Am. Chem. Soc.* 106:3623–3632.

51. Matsubara, T., F. Maseras, …, K. Morokuma. 1996. Application of the New "Integrated MO + MM" (IMOMM) Method to the Organometallic Reaction Pt(PR3)2 + H2 (R = H, Me, t-Bu, and Ph). *J. Phys. Chem.* 100:2573–2580.

52. Cummins, P. L., and J. E. Gready. 1997. Coupled semiempirical molecular orbital and molecular mechanics model (QM/MM) for organic molecules in aqueous solution. *J. Comput. Chem.* 18:1496–1512.

53. Cui, G. L., and W. T. Yang. 2011. Conical intersections in solution: formulation, algorithm, and implementation with combined quantum mechanics/molecular mechanics method. *J. Chem. Phys.* 134:204115.

54. Dokholyan, N. V., S. V. Buldyrev, …, E. I. Shakhnovich. 1998. Discrete molecular dynamics studies of the folding of a protein-like model. *Fold. Des.* 3:577–587.

55. Ding, F., W. Guo, …, J. E. Shea. 2005. Reconstruction of the src-SH3 protein domain transition state ensemble using multiscale molecular dynamics simulations. *J. Mol. Biol.* 350:1035–1050.

56. Dokholyan, N. V. 2006. Studies of folding and misfolding using simplified models. *Curr. Opin. Struct. Biol.* 16:79–85.

57. Ding, F., D. Tsao, …, N. V. Dokholyan. 2008. Ab initio folding of proteins with all-atom discrete molecular dynamics. *Structure.* 16:1010–1018.

58. Cammack, R. 1992. Iron-sulfur clusters in enzymes: themes and variations. *Adv. Inorg. Chem.* 38:281–322.

59. Gamiz-Hernandez, A. P., A. S. Galstyan, and E. W. Knapp. 2009. Understanding rubredoxin redox potentials: role of H-bonds on model complexes. *J. Chem. Theory Comput.* 5:2898–2908.

60. Gamiz-Hernandez, A. P., G. Kieseritzky, …, E. W. Knapp. 2011. Rubredoxin function: redox behavior from electrostatics. *J. Chem. Theory Comput.* 7:742–752.

61. Rutherford, K., I. Le Trong, …, W. W. Parson. 2008. Crystal structures of human 108V and 108M catechol O-methyltransferase. *J. Mol. Biol.* 380:120–130.

62. Hu, L., P. Söderhjelm, and U. Ryde. 2011. On the convergence of QM/MM energies. *J. Chem. Theory Comput.* 7:761–777.

63. Siegbahn, P. E., and M. R. Blomberg. 1999. Density functional theory of biologically relevant metal centers. *Annu. Rev. Phys. Chem.* 50:221–249.

64. Dirac, P. A. M. 1929. Quantum mechanics of many electron systems. *Proc. R. Soc. Lond. A Math. Phys. Sci.* 123:714–733.

65. Slater, J. C. 1951. A simplification of the Hartree-Fock method. *Phys. Rev.* 81:385–390.

66. Vosko, S. H., L. Wilk, and M. Nusair. 1980. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.* 80:1200–1211.

67. Becke, A. D. 1988. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A.* 38:3098–3100.

68. Perdew, J. P. 1986. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B Condens. Matter.* 33:8822–8824.

69. Arnim, M. V., and R. Ahlrichs. 1998. Performance of Parallel TURBOMOLE for density functional calculations. *J. Comput. Chem.* 19:1746–1757.

70. Sierka, M., A. Hogekamp, and R. Ahlrichs. 2003. Fast evaluation of the coulomb potential for electron densities using multipole accelerated resolution of identity approximation. *J. Chem. Phys.* 118:9136–9148.

71. Grimme, S. 2004. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comput. Chem.* 25:1463–1473.

72. Schafer, A., H. Horn, and R. Ahlrichs. 1992. Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *J. Chem. Phys.* 97:2571–2577.

73. Perdew, J. P., and Y. Wang. 1992. Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B Condens. Matter.* 45:13244–13249.

74. Tao, J., J. P. Perdew, …, G. E. Scuseria. 2003. Climbing the density functional ladder: nonempirical meta-generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.* 91:146401–146404.

75. Weigend, F., and R. Ahlrichs. 2005. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* 7:3297–3305.

76. Klamt, A., and G. Schüürmann. 1993. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans.* 2:799–805.

77. Ahlrichs, R., M. Bär, …, C. Kölmel. 1989. Electronic structure calculations on workstation computers: The program system Turbomole. *Chem. Phys. Lett.* 162:165–169.

78. Brooks, B. R., R. E. Bruccoleri, …, M. Karplus. 1983. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* 4:187–217.

79. Min, T., C. E. Ergenekan, …, C. Kang. 2001. Leucine 41 is a gate for water entry in the reduction of *Clostridium pasteurianum* rubredoxin. *Protein Sci.* 10:613–621.

80. Lin, I.-J., E. B. Gebel, …, J. L. Markley. 2003. Correlation between hydrogen bond lengths and reduction potentials in *Clostridium pasteurianum* rubredoxin. *J. Am. Chem. Soc.* 125:1464–1465.

81. Westler, W. M., I.-J. Lin, …, J. L. Markley. 2011. Hyperfine-shifted ^{13}C resonance assignment in an iron-sulfur protein with quantum chemical verification: aliphatic C-H$\cdots$S 3-center-4-electron interactions. *J. Am. Chem. Soc.* 133:1210–1216.

82. Eidsness, M. K., A. E. Burden, …, C. Kang. 1999. Modulation of the redox potential of the [Fe(SCys)(4)] site in rubredoxin by the orientation of a peptide dipole. *Biochemistry.* 38:14803–14809.

83. Marcus, R. A. 1956. On the theory of oxidation-reduction reactions involving electron transfer. I. *J. Chem. Phys.* 24:966–978.

84. Alexandrova, A. N., A. I. Boldyrev, …, L.-S. Wang. 2006. All-boron aromatic clusters as potential new inorganic ligands and building blocks in chemistry. *Coord. Chem. Rev.* 250:2811–2866 (and references therein).

85. Boldyrev, A. I., and L.-S. Wang. 2005. All-metal aromaticity and antiaromaticity. *Chem. Rev.* 105:3716–3757 (and references therein).

86. Dauter, Z., K. S. Wilson, …, J. Meyer. 1996. Zinc- and iron-rubredoxins from *Clostridium pasteurianum* at atomic resolution: a high-precision model of a ZnS_4 coordination unit in a protein. *Proc. Natl. Acad. Sci. USA.* 93:8836–8840.

87. Andersen, H. C. 1980. Molecular dynamics at constant pressure and/or temperature. *J. Chem. Phys.* 72:2384–2393.

88. Ramachandran, S., P. Kota, …, N. V. Dokholyan. 2011. Automated minimization of steric clashes in protein structures. *Proteins.* 79: 261–270.

89. Kabsch, W. 1976. A solution of the best rotation to relate two sets of vectors. *Acta Crystallogr. A.* 32:922–923.

90. Barton, G. J. 1993, 2002 "OC—A cluster analysis program", University of Dundee, Scotland, UK. http://www.compbio.dundee.ac.uk/downloads/oc.